DOCUMENT RESUME

ED 191 864 TM 800 475

AUTHOR Magnello, M. Eileen: Spies, Carl J.
TITLE The Use of Multidimensional Scaling with Statistical Concepts.
PUB DATE Apr 80
NOTE 31p.: Paper presented at the Annual Meeting of the American Educational Research Association (64th, Boston, MA, April 7-11, 1980). Computer produced matrices may be marginally legible.

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Concept Teaching: *Correlation: Goodness of Fit: Graduate Students: Higher Education: *Multidimensional Scaling: Pretests Posttests: *Statistics
IDENTIFIERS *Nonmetric Multidimensional Scaling: Stress Distributions

ABSTRACT

Shepard-Kruskal's nonmetric multidimensional scaling procedure (KYST) is used to discern those statistical concepts which are most confusing and/or ambiguous to graduate students, as well as to determine if a series of lectures in statistics will influence students' responses to more closely approximate those of their professors'. An instrument consisting of paired stimuli of nine elementary concepts measuring central tendency, variability, and correlation, were rated by 89 graduate students. Specifically, these concepts are: mean, median, mode, variance, standard deviation, semi-interquartile range, product-moment correlation, standard error of estimate, and standard error of the mean. The instrument was administered in two levels of statistic courses and pre- and posttest results were collected. KYST was used to analyze the data in one, two, and three dimensions. The 2- and 3-dimensional configurations did not lend themselves to interpretation. Introductory students showed the greatest change with the 1- dimensional configurations. As compared to the introductory pretest results, the posttest results for the introductory classes and the pre- and posttest results for the second-level classes yielded configurations more similar to the instructor's configuration. (Author/RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED191864

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

The Use of Multidimensional Scaling with Statistical Concepts

M. EILEEN MAGNELLO and CARL J. SPIES

Kent State University

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

M. Magnello

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Presented at the American Educational Research Association, Boston, Massachusetts, April, 1980

Requests for reprints should be sent to: M. Eileen Magnello
Department of Educational Foundations
406 White Hall
Kent State University
Kent, Ohio 44242

TM 800475

The use of Multidimensional Scaling with Statistical Concepts

M. Eileen Magnello and Carl J. Spies

Kent State University

Abstract

An instrument consisting of paired stimuli of nine elementary statistical concepts were rated by 89 graduate students in The College of Education during the Spring and Summer of 1979. The instrument was administered in two levels of statistic courses and pre- and post-test results were collected. Shepard-Kruskal's nonmetric multidimensional scaling procedure (KYST) was used to analyze the data in one, two, and three-dimensions. The 2- and 3-dimensional configurations did not lend themselves to interpretation. The 1-dimensional configurations yielded the clearest results, with the dimensions being labeled "complexity," for all but the pre-test configuration for students in Introductory Statistics. Introductory students showed the greatest change, from an uninterpretable pre-test configuration to the post-test configuration. As compared to the introductory pre-test results, the post-test results for the introductory classes and the pre- and post-test results for the second-level classes yielded configurations more similar to the instructor's configuration.

The use of Multidimensional Scaling with Statistical Concepts

M. Eileen Magnello and Carl J. Spies

Theoretical Framework

Psychological scaling methods can be viewed as procedures for constructing scales for the measurement of psychological attributes (Torgerson, 1958). Gustav Fechner, a German philosopher and physicist, was instrumental in developing scaling techniques for psychological attributes (Barid & Noma, 1978; Thurstone, 1927a; Torgerson, 1958). Fechner defines psychophysics as "an exact science of the functional relation of the dependency between body and mind" (Torgerson, 1958, p. 8). Fechner's development of psychophysics included both the measurement of sensory attributes and the quantification of perception, in order to correlate these psychological scales with physical measurement of the stimuli.

Louis L. Thurstone's (1927a) contention was that many of the "psychophysical" scaling methods could be used for accurate measurement of psychological attributes of stimuli which had no relevant measurable physical correlate. Consequently, Thurstone (1927b) developed the law of comparative judgment for data collected by Fechner's method of paired comparisons and showed it was possible to obtain internally consistent measurement for various psychological attributes, i.e., preference for various foods or gists. Since Thurstone's first paper

on the law of comparative judgment in 1927, a tremendous amount of work has appeared on the psychological scaling methods.

Scaling methods have been used to study the meaning of words, preference for sets of objects such as food, and goals of life. Scaling procedures have also been found useful in studies of discrimination learning, generalization, and personality (Torgerson, 1958). Psychologists have used multidimensional scaling to understand visual stimuli (i.e., colors and faces) and social entities such as personality traits (Rosenberg, Nelson, & Vivekanathan, 1968) and social stiuations. Educational researchers have used multidimensional scaling to study the structure of intelligence, or different test batteries, and of classroom environments.

Need for the Study

Multidimensional scaling could have heuristic value in education and science. It has been used to study the cognitive structures of subjects in psychology (Carroll & Wish, 1970, 1971; Dush & Gabriel, 1979; Horowitz, 1979), in education (Root, 1963; Traub & Hambleton, 1974), and science (Johnson, Cox, & Curran, 1970). The unifying purposes of multidimensional scaling are twofold; first to obtain a pattern or structure that may lie hidden in the matrix of empirical data, and secondly of representing that structure (in a form more accessible to the human eye) in a geometric model or picture (Shepard, 1972). Roger

Shepard (1963) interpreted the spatial configuration for data collected on the International Morse Code and named two dimensions. The first dimension was the number of components a letter or number had and the second dimension was the auditory signals in the Morse Code. This interpretation has heuristic value in that it can provide a visual aid to those learning or teaching the Morse Code. The teacher or learner of the Morse Code could know from this spatial configuration which codes (signals) could cause the most confusion in discriminating between two codes.

This same type of situation could be applied in many teaching situations, particularly in the field of statistics, where a novice is likely to experience much confusion and ambiguity. The need for the present study is to provide information on the effects of instruction in the field of statistics and to locate possible confusion students may manifest with statistical concepts.

The purposes of the study are twofold; the first concern is to discern those statistical concepts which are most confusing and/or ambiguous to graduate students. The other concern has potential educational value and is heuristic in nature; viz, the interaction of teaching and learning. Of particular interest is class intervention, that is, will the students' responses more closely approximate the professors' responses after a series of lectures in statistics than they did at the beginning of the term.

Methodology

The data in this study will be analyzed with KYST, Shepard-Kruskal's nonmetric multidimensional scaling procedure. The reasons for choosing KYST and the salient features of this method are as follows: (1) KYST assumes only that raw proximities (δ_{ij}) and distances (d_{ij}) are related by an unspecific monotonic function that is recovered as part of the scaling procedures; (2) it makes no linearity assumption and tends to produce fewer dimensions, thus facilitating representation and interpretation of a configuration; (3) it permits both Euclidean and non-Euclidean representations; and (4) it also operates easily with missing data (Subkoviak, 1975).

Subjects

The subjects in this study were Kent State University graduate students in the College of Education, and the College of Business. The sample was drawn from three graduate classes in statistics (2 classes from Introductory Statistics and the other from Inferential Statistics in the College of Education during the Spring and Summer of 1979. The final sample consisted of 42 students in Introductory Statistics and 47 students in Inferential Statistics.

Proximity Data

The proximity data formulated in this study were measures of central tendency, variability, and correlation. Specifically these concepts were

the; mean, median, mode, variance, standard deviation, semi-interquartile range, product-moment correlation, standard error of estimate, and standard error of the mean. The nine statistical concepts produced 36 possible combinations [n(n-1)/2], of paired stimuli which were then randomly ordered in the instrument.

Instrumentation and Procedures

The instrument was developed by the writers using purposive selection of the statistical concepts and then the concepts were paired and the pairings were randomly ordered in the instrument. The paired stimuli of statistical concepts were collected by the Method of Successive-Intervals. This method, formulated by Thurstone (1929) and refined by Torgenson (1958) required the students to rate all possible pairs on a similarity-dissimilarity scale of 1 to 9 where 1-Very Similar and 9=Very Dissimilar.

The instrument was administered to all three graduate classes in statistics in the College of Education. The instrument was distributed twice, first at the beginning of the term and secondly near the end of the term. The professor of the statistic classes completed the instrument in order that necessary comparisons could be made. The students were allocated as much time as they needed to complete the instrument.

Statistical Background and Procedures

The responses to the instrument from the students and the instructor were analyzed with KYST (Kruskal, Young, Shepard, and Torgerson), a

nonmetric multidimensional scaling computer program, written by Kruskal, Young, and Seery (1975). KYST places N points in a space of dimension so as to minimize "stress" which measures the "goodness-of-fit" between the configuration of points and the data. In order to obtain the "stress" value the data is iterated and KYST uses the iterative numerical method of gradients (the method of steepest descent). Stress is a value which represents the variance unaccounted for by a given dimensional solution. This value is analagous to the standard error of estimate in a bivariate regression.

After the data has been iterated and the stress value found, the number of dimensions are determined. Kruskal and Wish (1978) have stated the following considerations in choosing an appropriate dimensionality; goodness-of-fit (stress), interpretability, ease of use, and stability. The trade relation with parsimony and interpretability on the one hand and goodness-of-fit on the other, is that as the number of dimensions increases goodness-of-fit always improves, but at the expense of a more interpretable and parsimonious solution.

The final step in multidimensional scaling is the interpretation of the resultant dimensions. This step tends to be subjective on the investigator's part. The dimensionality and in particular the orientation of the axes (i.e., the direction in which these lines project in space) with respect to the stimulus location offer much assistance in interpreting

the underlying structure. Torgerson (1965) asserts that the best guide to interpreting the location and axes is to develop a good a priori theory of the underlying structure. Baird and Noma (1978) espouse Torgerson's view by stating that if a theory explicitly states that results should fall into two dimensions, then such a stricture should be followed even if stress were relatively high.

A 9 x 9 proximity matrix was generated from the 36 pair-wise comparisons which consisted of the mean similarity judgments of these comparisons. Since each group had rated the form twice (pre- and post) there were two proximity matrices per group. The instructor's matrix then produced the fifth proximity matrix and these matrices were subsequently analyzed using KYST. A monotonic regression was performed on the data in order that the data could be analyzed nonmetrically in 1,- 2,- and 3- Dimensions. Young (1970) and Sherman (1972) have suggested the minimum number of objects to use in order to obtain a satisfactory 1, 2, or 3- dimensional solution is 5, 9, or 13 objects respectively.

Results

The results of this study revealed that the 2- and 3- dimensional solutions were uninterpretable for both groups and the instructor. Although the stress values for the 2- and 3- dimensional solutions were lower than the stress values for the 1- dimensional solution, it did not lend itself to interpretability and consequently will not be presented

here. Since only 9 stimulus objects were used then it is not surprising that a 3- dimensional solution was uninterpretable.

The stress values for the 3- dimensional solutions are as follows: students from introductory statistics pre-test, stress = .174; post-test, stress = .168; students from inferential statistics pre-test, stress = .167; post-test stress = .172; and the instructors stress value = .017. The stress values for the 2- dimensional solutions are as follows: students from introductory statistics pre-test, stress = .271; post-test, stress = .247; students from inferential statistics pre-test, stress = .255; post-test stress = .263; and the instructor's stress value = .068. Although the 1- dimensional stress values were quite high, thus representing a "poor" goodness-of-fit, they were interpretable. Kruskal (1975) and Shepard (1972) maintain that even if the stress value is high (for a smaller dimension) but the solution in that dimension is interpretable, then the interpretable and more parsimonious solution should be chosen, hence, the "trade-relation."

The stress values in a 1- dimensional solution for the students in the Introductory Statistic class are as follows: for the pre-test, the 1- dimensional stress = .455; for the post-test, the 1- dimensional stress = .394. The resulting dimensional saliences for these solutions are given in Table 1. The dimension for the pre-configuration was uninterpretable and the dimension for the post-configuration was labeled "complexity."

The concepts in the table are presented in the order in which they ascended in the configurations, i.e., for the pre-test the mode was the first concept and variance was last, and for the post-test the mode was the first concept and standard error of estimate was the last concept.

Table 1

Dimensional Saliences for the Students in Introductory Statistics of the 9 Statistical Concepts in 1- Dimension

Concept	Loadings	Concept	Loading
Pre-Test		Post-Test	
Mode	-1.526	Mode	-1.590
Correlation (r)	-1.221	Semi-Interquartile Range Q	-1.075
Median	-0.766	Median	-0.846
Standard error of the Mean (Sx)	-0.329	Mean	-0.432
Mean	-0.068	Standard Deviation (S)	0.072
Standard Deviation (S)	0.586	Variance (S^2)	0.412
Semi-Interquartile Pange Q	0.592	Standard Error at the Mean (Sx)	0.752
Standard error of Estimate (Sy.x)	1.199	Correlation (r)	1.262
Variance (S^2)	1.533	Standard Error of Estimate (Sy.x)	1.441

The spatial configurations generated from these solutions can be seen in Figure 1 for the pre-test and in Figure 2 for the post-test.

The stress values in a 1- dimensional solution for the Inferential Statistic class are as follows: for the pre-test, stress = .430, and for the post-test, stress = .454. The resulting dimensional saliences for these solutions are given in Table 2. The dimensions for the pre- and post-configurations were also labeled "complexity." The concepts for in the table are presented in the order in which they ascended in the configuration.

Table 2

Dimensional Saliences for the student in Inferential Statistics of the 9 Statistical Concepts in 1- Dimension

Concepts	Loadings	Concepts	Loadings
Pre-Test		Post-Test	
Standard Error of Estimate (Sy.x)	-1.543	Mode	-1.538
Standard Error of the Mean (Sx)	-1.106	Mean ($\bar{X}$)	-1.115
Variance	-0.768	Median	-0.915
Standard Deviation (s)	-0.436	Semi-Interquartile Range Q	-0.410

Concepts	Loadings	Concepts	Loadings
Pre-Test		Post-Test	
Correlation (r)	-0.092	Correlation (r)	0.046
Semi-Interquartile Range Q	0.354	Standard Error of the Mean ($S_{\bar{x}}$)	0.518
Median	0.967	Variance (S^2)	0.839
Mean ($\bar{X}$)	1.086	Standard Deviation (S)	1.064
Mode	1.538	Standard Error of Estimate (Sy.x)	1.510

The spatial configurations generated from these solutions can be seen in Figure 4 for the pre-test and in Figure 5 for the post-test.

The stress value in a 1- dimensional solution for the instructor was .227. The resulting dimensional saliences for the solution is given in Table 3, and are presented in the order in which they ascended in the configuration. This dimension was labeled "complexity."

Table 3

Dimensional Saliences for the Instructor of 9 Statistical Concepts in 1- Dimension

Concepts	Loadings
Correlation (r)	-1.564
Standard Error of Estimate (Sy.x)	-1.013
Variance (S^2)	-0,634

Standard Deviation (s)	-0.440
Standard Error of the Mean (Sx)	-0.160
Mean ($\overline{X}$)	0.246
Semi-Interquartile Range (Q)	0.629
Median	1.204
Mode	1.732

The spatial configurations generated from these solutions can be seen in Figure 3.

DISCUSSION

The configurations for the Introductory Statistics students reflected the greatest change in their ordering of these concepts. Most of the students in this group had little or no prior knowledge of statistics at the beginning of the course. Those students who had some prior knowledge of statistics generally had an undergraduate class in statistics. There was little change from pre- to post-configuration for the students in Inferential Statistics. This is not surprising since all of these students had Introductory Statistics prior to this class. Of particular concern to the writers was the comparisons of the Inferential pre-configuration to the Introductory Statistics post-configuration. These configurations were similar with the exception of (r) the Pearson correlation

coefficient, the placement of r can be partially explained by the fact that some of these students had other instructors for Introductory Statistics. Since the two configurations were similar in other respects, perhaps this reflects (for students in Inferential Statistics) some retention of knowledge from Introductory Statistics.

The pre-configuration (for students in Introductory Statistics) reflected a certain degree of randomness along with some degree of knowledge of being able to differentiate between measures of central tendency and measures of variability. However, the post-configuration revealed that these students could see other existing relationships among various concepts [i.e., between r (correlation) and Sy.x (standard error of estimate) and between the median and Q (semi-interquartile range)] that were not revealed in the pre-configuration.

The ordering of the 1- dimensional solution (from the post-configuration for the Introductory Statistic students) was of particular interest. The lower-left-hand corner of this configuration started with the mode and ascended to Sy.x (standard error of estimate), from the easiest to the more difficult, computationally and conceptually. This dimension was appropriately labeled "complexity" as it appears to reflect their understanding of these concepts which may also be relative to time. The ordering of the concepts was similar to the way in which these concepts were introduced in class, of the concepts used in study.

The instructor first introduced central tendency, then variability, and finally the Pearson correlation along with Sy.x (standard error of estimate). As was anticipated, the post-configurations for these students does resemble the ordering of the instructor's configuration.

Their configuration seems to indicate that the course of instruction had an effect on the students learning not just in the ordering of these concepts, but also by the placement of these concepts along the axis which suggested a systematic ordering of the concepts.

The analyses in this study provided a "stepping-stone" for further inspection of the data. Further analyses will be done using INDSCAL, a multidimensional scaling computer program which yields individual dimensional saliences and configurations (Carroll & Chang, 1970). These configurations will allow for individual student comparisons with the professor's configurations (perhaps yielding other dimensions) and these individual saliences can also be used to obtain the correlation between their salient loadings and the final grade these students received in their statistics class. Since this study does reflect change with regards to learning, it appears to be worthwhile to further investigate this by making comparisons across other disciplines (i.e., Psychology, Sociology, and/or Business) with graduate students enrolled in statistic classes. These comparisons may yield other dimensions or dimensions that can be interpreted differently and this knowledge may provide further insight into their use and perception of statistics.

For Configuration Plot from Computer Program

A = S^2 (variance)

B = r (Pearson product-moment correlation)

C = $S_{\bar{x}}$ (standard error of the Mean)

D = $\overline{X}$ (Mean)

E = S (Standard Deviation)

F = Sy.x (Standard Error of Estimate)

G = Q (Semi-Interquartile Range)

H = Median

I = Mode

Figure 1

One-Dimensional Spatial Configuration of
Nine Elementary Statistical Concepts
for Students in Introductory Statistics
(Pre-Test)

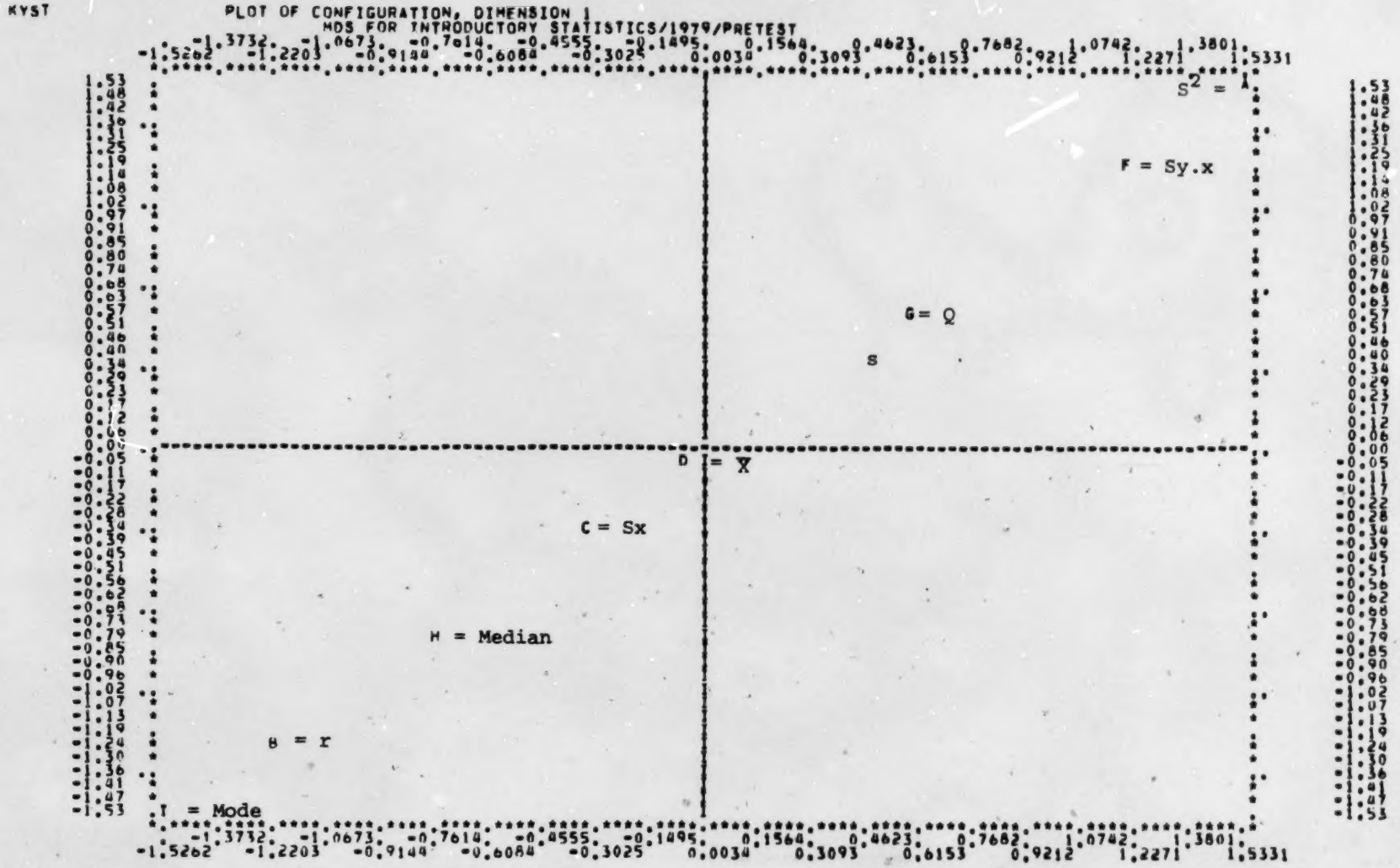


Figure 2

One-Dimensional Spatial Configuration of
Nine Elementary Statistical Concepts
for Students in Introductory Statistics
(Post-Test)

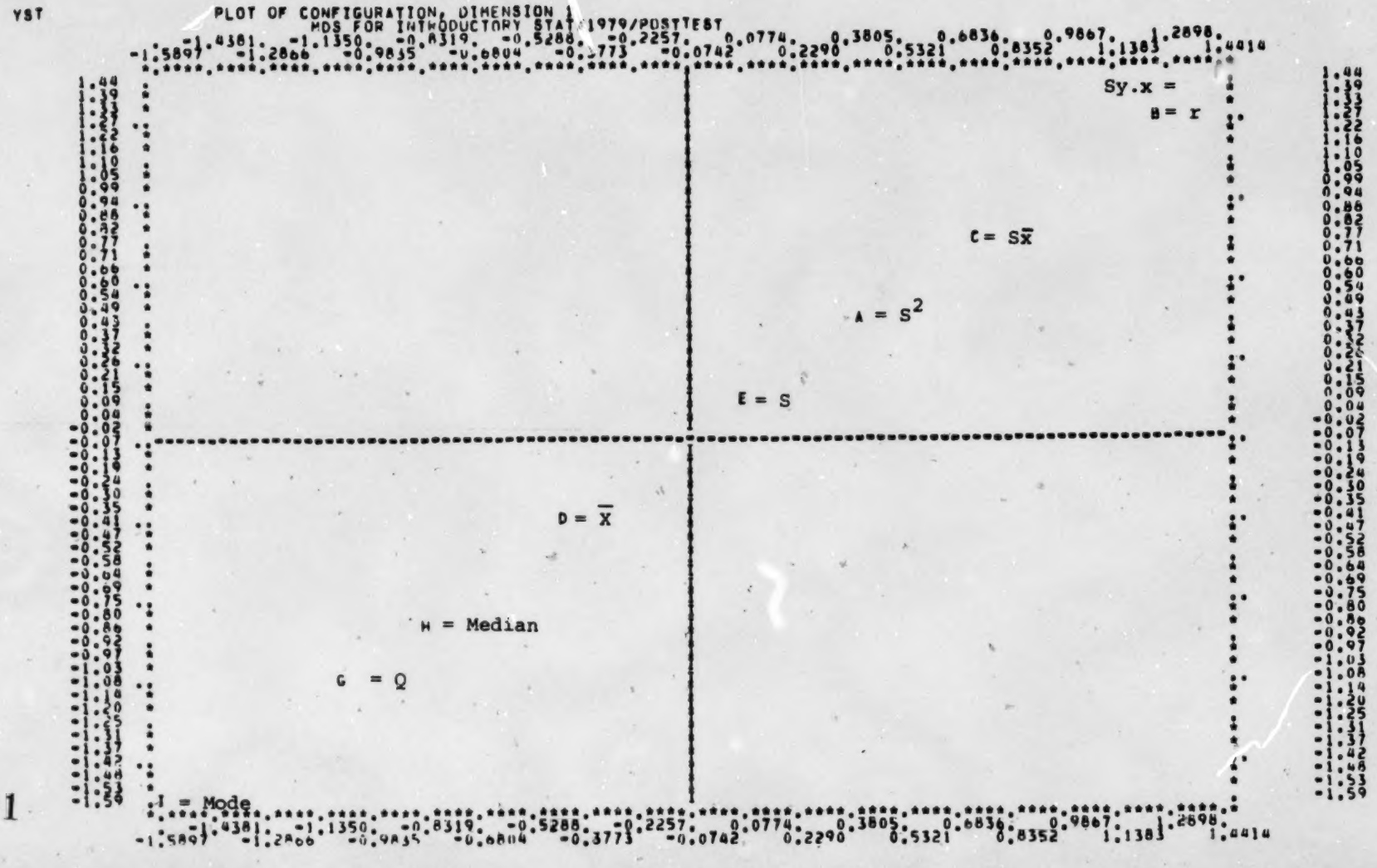


Figure 3

One-Dimensional Spatial Configuration of
Nine Elementary Statistical Concepts
for Students in Introductory Statistics
for the Instructor

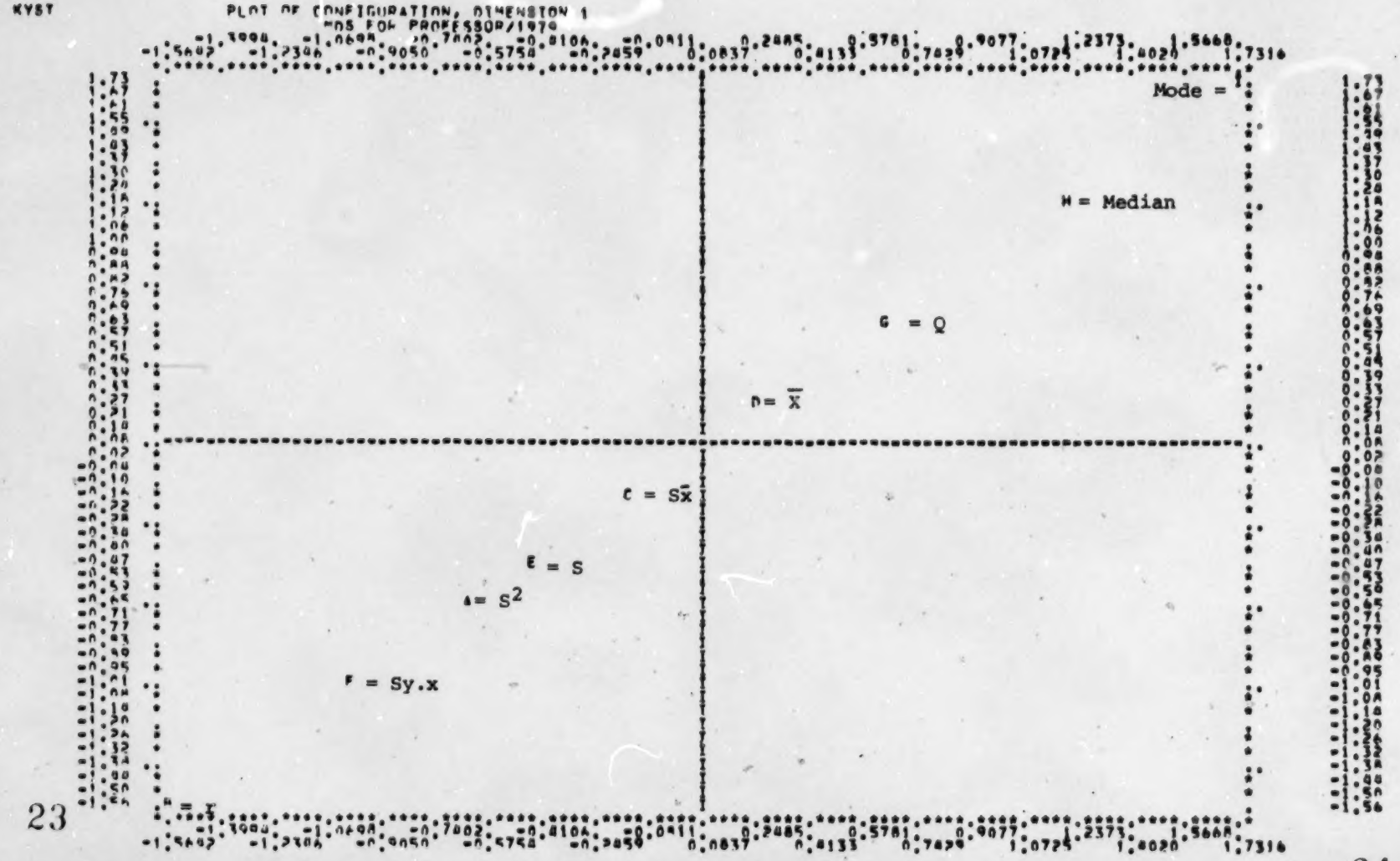


Figure 4

One-Dimensional Spatial Configuration of
Nine Elementary Statistical Concepts
for Students in Inferential Statistics
(Pre-Test)

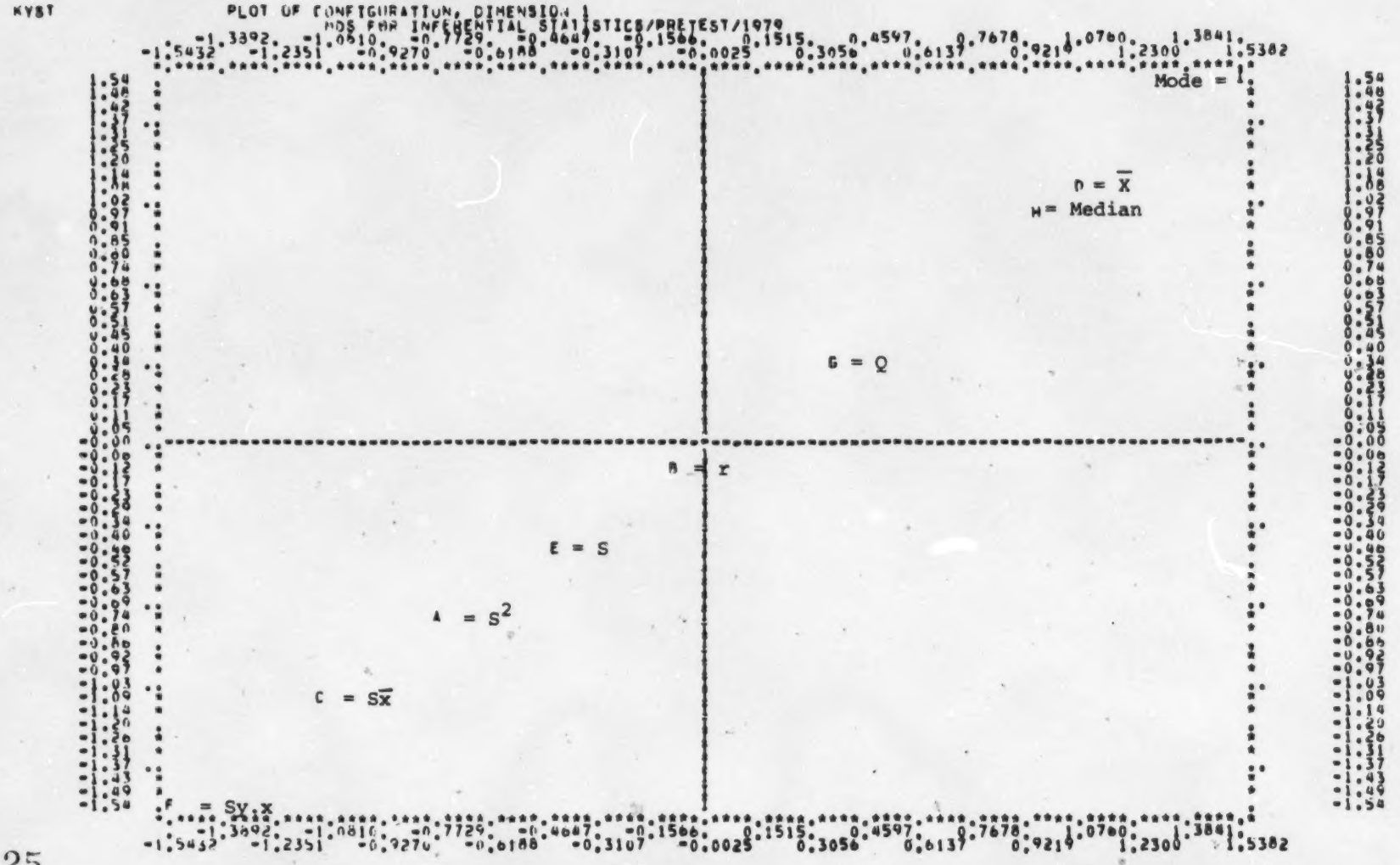
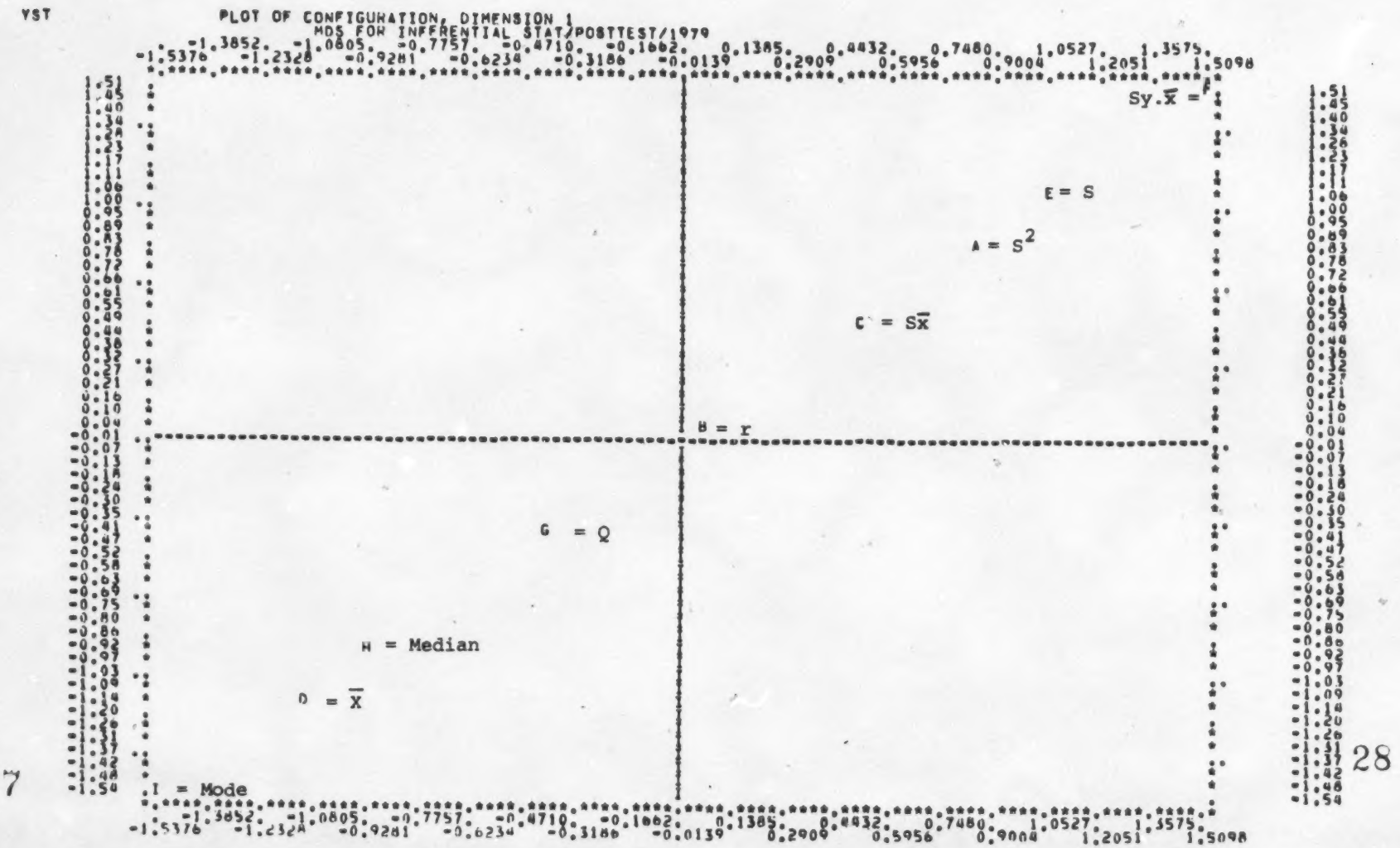


Figure 5

One-Dimensional Spatial Configuration of
Nine Elementary Statistical Concepts
for Students in Inferential Statistics
(Post-Test)



REFERENCES

Baird, J. C. & Noma, E. Fundamentals of scaling and psychophysics. New York: John Wiley & Sons, Inc., 1978.

Carroll, J. D., & Chang, J. J. Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" Decomposition. Psychometrika, 1970, 35(3), 283-319.

Carroll, J. D., & Wish, M. Multidimensional scaling of individual differences in perception and judgment. (To appear in Bell Laboratories Record), Unpublished Manuscript, Bell Laboratories, 1970.

Carroll, J. D., & Wish, M. Measuring preference and perception with multidimensional models. Bell Laboratories, 1971, 147-154.

Dush, D. & Gabriel, R. An analysis of the multidimensionality of inkbolt perception. Paper presented to the 87th Annual Convention of the American Psychological Association, (Division 5), New York City, 1979.

Horowitz, L. M. On the cognitive structure of interpersonal problems treated in psychotherapy. Journal of Consulting and Clinical Psychology, 1979, 47(1), 5-15.

Johnson, P. E., Cox, D. C. & Curran, T. E. Psychological reality of physical concepts. Psychonomic Science, 1970, 19, 245-247.

Kruskal, J. B. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. Psychometrika, 1964, 29(1), 1-27.

Kruskal, J. B. Nonmetric multidimensional scaling: A numerical method. Psychometrika, 1964, 29(2), 115-129 (b)

Kruskal, J. B., Young, F. W., & Seery, J. B. How to use KYST, A very flexible program to do multidimensional scaling and unfolding. Unpublished paper, Bell Laboratories, 1973.

Kruskal, J. B., & Wish, M. Multidimensional Scaling. Sage University Paper series on Quantitative Applications in the Social Series, 07-011. Beverly Hills and London: Sage Publications, 1978.

Root, R. T. A multidimensional scaling analysis of meaning. Dissertation Abstracts, 1963, 23, 1979.

Rosenberg, S., Nelson, C., & Vivekananthan, P. S. A multidimensional approach to the structure of personality impressions. Journal of Personality and Social Psychology. 1968, 9(4), 283-294.

Sherman, C. R. & Young, F. W. Nonmetric multidimensional scaling: A Monte Carlo study. Reprinted from the Proceedings, 76th Annual Convention, APA, 1968.

Shepard, R. N. Analysis of proximities as a technique for the study of information processing in man, Human Factors, 1963, 5, 33-38.

Shepard, R. N. Purposes of multidimensional scaling. In R. D. Shepard, A. K. Romney & S. B. Nerlove (Eds.), Multidimensional Scaling (Vol. 1) New York: Seminar Press, 1972.

Subkoviak, M. J. The use of multidimensional scaling in educational research. Review of Educational Research, 1975, 45(3), 387-423.

Thurstone, L. L. A metal unit of measurement. Psychological Review, 1927, 34, 415-423. (a)

Thurstone, L. L. Three psychological laws. Psychological Review, 1927, 34, 424-432 (b)

Thurstone, L. L. Fechner's law and the method of equal-appearing intervals. Journal of Experimental Psychology, 1929, 12, 214-224.

Torgerson, W. S. Theory and methods of scaling. New York: John Wiley & Sons Inc., 1958.

Traub, R. E., & Hambleton, R. F. The effect of instruction on the cognitive structure of statistical and psychometric concepts. Canadian Journal of Behavioral Sciences, 1974, 6(1), 30-43.

Young, F. W. Nonmetric multidimensional scaling: Recovery of metric information. Psychometrika, 1970, 35(4), 455-473.